Assessing Risk Levels of Verification, Validation, and Accreditation of Models and Simulations

James N. Elele, Ph.D

U.S. Naval Air Warfare Center (NAVAIR), Patuxent River, Maryland

This article presents a new methodology for assessing the risks associated with the level of verification, validation, and accreditation (VV&A) of a given model and/or simulation when used in support of major decisions. As stated by Department of Defense (DoD) Instruction 5000.61 (DOD, 2003),[1] "It is DoD policy that: models and simulations (M&S) used to support major DoD decision-making organizations and processes shall be accredited for that specific purpose by the DoD component M&S application sponsor." Instruction 5000.61 applies to "All models and simulations developed, used, or managed by the DoD components after the effective date of this instruction." The requirements cited above have set the need for VV&A of M&S at the forefront of concerns for DoD and DoD-Component acquisition personnel. When an acquisition program involves a large number of models, cost associated with VV&A can become enormous. There is a need therefore to have a systematic approach for assessing and prioritizing the risks associated with the level to which individual models have been verified, validated, and accredited. To provide decision-makers with a judicious way for determining the risks associated with using a given M&S, and the extent to which VV&A work will be needed to meet these requirements, a methodology is developed for assessing the risks associated with the level of VV&A of a given M&S when used to support decision-making. This approach parallels the formal DoD Risk Assessment procedure, but with application to the use of M&S, as it relates to VV&A.

Key words: Accreditation; consequence of error; likelihood of error; modeling & simulation credibility; risk assessment; validation; verification.

Modeling & Simulation (M&S) have become an integral part of the defense acquisition process in the United States. The value of M&S in various aspects of the defense acquisition process will continue to increase as advancement in computer technology and new modeling techniques emerge. Since resources are limited, there is a need for a logical way to guide and prioritize the investment in M&S while making sure that models are credible, particularly when such M&S are used to support very important decisions. Moreover, M&S saves money by reducing the cost associated with system design, development, and testing. But, in addition to cost savings, there is a less emphasized but very valuable role of M&S in military acquisitions: M&S allow system developers to reduce and/or avoid risks associated with human involvement when investigating the boundaries of applicability/survivability of

systems. With M&S playing such critical roles in war fighter systems development, there is a critical need to devise verification, validation, and accreditation (VV&A) approaches that use available resources efficiently without compromising the needed credibility of the M&S.

Formal VV&A can be very expensive. How much (and what kind of) evidence is required for establishing confidence in, and reaching an accreditation decision for, a particular M&S determines the amount of resources required. For existing M&S (e.g., legacy M&S that have been used in the past without formal VV&A), current methodologies provide the users with little, if any, way to assess the risks associated with accepting and using such a model or a simulation when its credibility is in question; and the latter is often the case even though the model may have a history of extensive use. What is usually done is to accept and use the M&S with or without VV&A. But when the model

		Form Approved OMB No. 0704-0188

Report Documentation Page

1. REPORT DATE **2008**	2. REPORT TYPE	3. DATES COVERED **00-00-2008 to 00-00-2008**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Assessing Risk Levels of Verification, Validation, and Accreditation of Models and Simulations	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Naval Air Warfare Center (NAVAIR),47123 Buse Rd, B2272 Unit IPT,Patuxent River,MD,20670-1547**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**11**	

result deviates seriously from what is expected, the user may have to start from scratch and reengineer the model development process in order to generate the data and artifacts needed for establishing credibility.

The methodology we present here uses information from various aspects of the development of the M&S — its VV&A history, knowledge from subject matter experts (SMEs), the role the M&S will play in the decision-making, and the consequence if the M&S is wrong — to arrive at a categorized risk level (high, medium, or low risk) expected with accepting such an M&S for an intended use. Determining this risk allows the user to choose to mitigate the identified risk by conducting formal VV&A, or to accept the risk, accredit the M&S with limitations, and continue to use it as is. This approach becomes very useful when a large number of models and simulations are needed to support an acquisition program, so there is a need to prioritize and allocate M&S-related funds and resources efficiently.

The issue is credibility of M&S for an intended use

The goal of VV&A is to generate, maintain, gather, and apply information about a given M&S to support the decision to use that model. The problem has been that VV&A was frequently discussed rather than practiced; and when implemented, it was usually done as an afterthought. One thing no one argues against is that it is very important for M&S to be credible and suitable for the specific intended use. But credibility and appropriateness of M&S for intended use are best established through the VV&A process. Thus M&S users need to recognize that VV&A is necessary for risk reduction and critical for establishing credibility. Because VV&A is frequently tailored to specific needs, a formal definition of VV&A is presented. The definitions given below are congruent to those of DoD, Navy, and other Services; and are particularly suited for practical applications and the pedagogical purposes we pursue here.

- **Verification:** Verification is the process of determining that a model implementation and its associated data accurately represent the developer's conceptual description and specifications. The practical question answered by verification is "Is the model relatively error free, and does it do what the originator intended?"
- **Validation:** Validation is the process of determining the degree to which a model and its associated data are an accurate representation of the real world from the perspective of the intended uses of the model. Note that M&S validation is not the same as software validation.

The practical question answered by verification is "Do model results match real world data well enough for the user's needs?"

- **Accreditation:** Accreditation is the official determination and certification that a model, simulation, or a federation of models and simulations and its associated data are acceptable for use for a specific purpose. The practical questions answered by accreditation are "Does the accreditation authority have adequate documented evidence to be confident that a model and its input data are credible for a particular use; and is there enough documented information to show that this M&S is fit for this purpose?"

It is reemphasized here that VV&A, in principle, is a process for reducing risk; in that sense VV&A provides a way for establishing whether a particular M&S and its input data are suitable and credible for a particular use. According to *M&S VV&A Implementation Handbook*, Volume 1,[2] risk management as related to VV&A is simply answering the question "What is the risk of using a model compared to using the real world system or methodology?" VV&A provides the answer to this question through processes and procedures designed to mitigate the risks while providing objective evidence as needed to establish the credibility of the M&S.

Statistical basis for risk-based VV&A

Risk-based VV&A is based on the statistical principle of hypothesis testing. Given any M&S, VV&A has become the standard accepted way to answer the question "Is this M&S credible for the intended use?" In statistical language, this problem is stated as testing the "Null Hypothesis," which asserts **Ho: *The M&S is credible for the intended use.*** (One usually must also state an "Alternate Hypothesis"; e.g., **Ha: *The M&S is not credible for the intended use.***) To complete such a test, one would set up an experiment, collect data, and test the above hypotheses for acceptance or rejection. The data so collected is considered as sample points taken from a known or assumed probability distribution for the events occurring in the experiment. Since this type of testing is based on sampled data, there is always the possibility of making errors; and the two categories of possible errors that result are, namely, Type I and type II errors:

- A **Type I error** denoted by α; occurs when one rejects H_0 when it is true.
- A **Type II error** denoted by β; occurs when one accepts H_0 when it is false.

The individual conducting the test must then specify the maximum allowable probability of making a Type I error (α, called the level of significance). The problem is that, in general, the experimenter is not able to

control the probability of making a Type II error. From the M&S perspective, making Type II errors is equivalent to accepting that M&S is credible for the intended use when it is not. Because one is not able to control this type of error in the experiment, there is not much one can do (in theory) to control the probability of Type II error where M&S is concerned. In practice, however, the acceptance or rejection of a model for an intended use is only done through formal VV&A assessment of the M&S under consideration.

Typically, statisticians try to avoid the risk of making a Type II error by using an expression such as "**cannot reject H_0**" based on the evidence/data, and by actually avoiding to say, "**accept H_0**" whenever possible An alternative approach uses the sampled data to compute a statistic called a **P value**, (which is the probability of obtaining a sample result that is at least as unlikely as the data that was observed). If the **P value** is less than the level of significance α, the tester rejects the null hypothesis, and accepts it otherwise.

It is the above statistical test structure that gives rise to the risk associated with using M&S and actually forms the theoretical basis for the risk-based VV&A approach.

Obviously, once a decision to use M&S has been made, one automatically becomes subjected to making these types of errors. Type I error is made when a user rejects a model or simulation that is credible for the intended use. But making Type I error is not considered very critical in real life (though this type of error can become very important from a cost-savings perspective). This fact is because the main loss for a user who makes an M&S-related Type I error (i.e. rejecting to use an M&S as not credible when it really is) is opportunity lost/cost. However, if a user decides to apply M&S, that user has accepted the chance of committing Type II error, and the consequences of such an error can be substantial. Thus it is emphasized that the principal risk associated with using M&S in decision-making is defined by Type II error. Moreover, the likelihood of making such an error and the associated consequence determine the risk associated with using that M&S in decision-making. VV&A therefore provide the accepted practical method for reducing the risk associated with using the M&S and establishing confidence in the M&S. The consequences of the model being wrong, and the level of risk one can accept, drive the amount of effort required to establish an acceptable credibility level for the M&S.

Definition of risk in the risk-based VV&A paradigm

The Defense Modeling and Simulation Office (DMSO) defines *risk* as "the potential realization of undesirable consequences from hazards arising for a possible event".[3] Of course, risk is a factor in any kind of decision-making in which imperfect evidence is used to help make the decision. Though VV&A is oftentimes more concerned with operational risk (i.e., the risk associated with using the M&S), the approach presented here can also be used, and has been used, to assess the risks associated with M&S development (i.e., the risk associated with completing the M&S development on time and within budget). This practical approach allows the System Safety community to define risk as the product of the likelihood of error and the consequences associated with such an error.

DEFINITION: Risk = (Likelihood of Error) * (Consequence of Error)

In this definition, *likelihood of error* is the probability that the M&S and/or its input data are incorrect or inappropriate for the intended use. The *consequence* is defined as the impact if the M&S output is wrong and the user accepts and uses it as correct. Thus to reduce risk, one either has to reduce the likelihood that something will go wrong or reduce the severity of the consequence(s) or impact/effect that will result when something goes wrong.

The consequence or impact when a model is wrong depends on the role the M&S will play and how important this role is in the decision-making process. One can choose not to use M&S (for example, by choosing to only accept actual physical measurements in decision-making), thus avoiding the risk associated with M&S use. Alternatively, one can choose to reduce the risk associated with using M&S by limiting the role M&S will play in the decision process. However, the reality is that DoD, the Navy, and other Services have mandated the use of M&S (ubiquitously) in the acquisition process, so that what is really needed is a method that will give M&S users the ability to assess the risks associated with the use of M&S, while providing them with the capability to mitigate such risks. This is what the current method tries to do.

Risk associated with M&S may be viewed differently from different managerial points of view. For example, an M&S developer/program manager may be more interested in the risk associated with the delivery of a simulation on time (schedule) and on budget (a form of development risk). Verification and Validation (V&V) is concerned with mitigating development risk by collecting objective evidence as required to demonstrate the capability, relative freedom from defects, level of fidelity, and the accuracy of representation of reality. From this perspective, part of what V&V is concerned with is providing the artifacts needed for accreditation. Providing (or requiring more) evidence

needed to certify usability, limitations, and fitness for intended use is the goal of accreditation. Consequently, accreditation is more concerned with exposing operational risk associated with M&S use.

Problem

As precise as the above definition of risk is, it has problems. To determine the risk, one must be able to determine both the likelihood of error and its consequence precisely, and then be able to multiply these two values to compute the value of the risk. But for M&S, one is not usually able to determine the likelihood of error precisely. Even when one is able to assess the consequence(s) of such an error, the data required to determine the likelihood of error precisely is usually lacking (recall that statisticians do not have the ability to control Type II errors). How then can one multiply something not determined? One cannot!

What is done is to classify each of these parameters (risk, likelihood, and consequence) into discrete scales or levels sometimes called baskets or bins: e.g., high, medium, and low for risk; frequent, probable, occasional, remote, and improbable for likelihood; and some form of numerical rankings for consequence. Doing this allows the user to apply the above scales in a systematic way, in an attempt to get a handle on the risks associated with using M&S to support decision-making. Understanding how this works forms the practical basis for the risk-based VV&A procedure, and this approach also makes it possible for an M&S user to determine how much effort to put into VV&A, the appropriate and extent of the review process, the level of independence in V&V review, as well as the appropriate level of accreditation authority. More details are provided on how this methodology can be used to support the accreditation process in a companion article.

In practice

We now turn attention to the practical applications, "The How To" of this method. This is necessary to allow more users to use this method, and because the more people there are who understand and use this methodology the better it will become. The method is broad enough that it can be adapted for specific applications. To start, some of the specific tools needed to make the method work are identified.

Where to begin

When faced with a VV&A question about a given model, the ease or difficulty of the validation process depends on whether a version of the M&S system currently exists or if it is being built for the first time. Designing the VV&A processes into a new model while it is being built is the better and preferred approach. In any case, the first task is usually to determine the scope of the effort necessary to accomplish the needed V&V, and the attendant accreditation for the specific intended use.

This problem requires that a cost-effective VV&A plan be devised. Such a plan needs to include a well-articulated Intended Use Statement (IUS). The IUS will help in defining the role the M&S will play in the decision process. The IUS may have already been written as part of the M&S requirement document; however, the role of the M&S in decision-making may best be determined through the knowledge of such things as available technology, planned activities, cost and schedule, and program management priorities. This requires one to answer such questions as "Will M&S be the only tool that will be used to generate the data needed for a major decision, or will other activities such as laboratory/actual test and evaluation (T&E) data be used with data from M&S to support the decision?" At a minimum, one must answer the question, "Will this M&S be used to support a minor, medium, or major acquisition decision?"

Sample scales and tables

This procedure will require the implementer to develop or adapt/adopt various scales and tables for use as classification buckets. The Battlespace Verification, Validation, and Accreditation Support Office (BVVASO)/Joint Accreditation Support Activity (JASA) has developed a series of tables and scales based on years of practical implementation of VV&A procedures for various acquisition programs (BVVASO, October, 1997).[4] Our experience from working with acquisition programs indicates that some of these tables would be invaluable, at least, as a starting point for practical implementation of the methodology. We provide a list of the tables and scales that a user might need when implementing the method in subsequent sections.

Table 1 gives a sample confidence/likelihood scale that is based on BVVASO's experience and guidelines in *DMSO VV&A Recommended Practice Guide (RPG)*.[5] It is recommended that one level be included in this table for either low or unknown level of confidence to allow for a minimal effort option, and to cover emergency or low consequence situations.

Table 2 provides (from BVVASO experience) a possible way to scale likelihood of error and/or confidence level. The table summarizes what has been found to be the information necessary to support an accreditation assessment.

- The higher the likelihood of error (or the lower the confidence in the M&S), the more need there

is for rigorous oversight and review through documentation, code and logic verification, configuration management, review of model development approach, data and code validation, to drive down likelihood of error.

- As likelihood of error goes down, confidence in model results goes up. Using the table below, one is able to assess the likelihood of error in an existing model or even determine what further evidence or level of accreditation is needed.

Table 2 is based on BVVASO's rules which have been adopted by the *DMSO VV&A RPG*[5]. See "Role of Accreditation Agent in VV&A of Legacy Simulations" for more details (available at www.vva.dmso. mil). *Table 3* is an example scheme quantifying likelihood.

Table 4 gives a sample scale for identifying and categorizing the importance of the role M&S will play in the decision-making. For example if M&S will be the only tool for making a decision, as is often the case in survivability studies, then the impact of this decision being wrong can be catastrophic, hence it is important to consider the risk associated with the level of VV&A of any such M&S.

Table 5 presents a scale for categorizing the level of importance of a decision based on the intended use of any model or simulation. This table provides a scale for accounting for the risk associated with errors resulting from M&S influence on such decisions. *Table 6* gives an example of the level of consequence on performance (cost, schedule, and technical quality) if the decision is a poor one. Criticality measure is determined from the level of reliance on M&S and importance of the decision. Criticality measure drives the nature and amount of information and effort applied to VV&A of this model *(Figure 1)*.

Table 7 gives a very simplified example of consequence scale with four broadly defined levels. Typically, consequence scales are defined with five levels for ease of application. This is because the standard risk chart has normally been defined with five levels of consequence. In practice, the choice of levels may best be determined by the circumstance being addressed by the user.

Table 8 gives a more advanced example of consequence scale with four broadly defined levels (Catastrophic, Critical, Marginal, and Negligible). Typically, Consequence scales are defined with five levels for ease of application. This is because the standard risk chart has normally been defined with five levels of consequence. In practice, the choice of levels may best be determined by the circumstance being addressed by the user.

Rules for associating/combining scales and tables

Once the scales and tables have been assembled, the user will also need to use combination/association rules to determine the risk associated with M&S. For example, the following rules were used in practical applications:

- **Rule 1: Value of "role (or level of reliance) of M&S (gray color)" in decision-making & value of "the importance of decision (gray color)" →** **"consequence of model error (green color)."** Determine the level of consequence if the model is in error by associating the role of M&S in decision-making and the level of importance of decision.
- **Rule 2: Value of the "level of reliance (gray color)" & the "confidence level (gray color)" → "likelihood of model error (gray color)."** Then

Importance of Decision	Level of Reliance on M&S			
	4	3	2	1
4	4	4 or 3	3 or 2	2
3	3	3	2	2 or 1
2	2	2	2	1
1	1	1	1	1

Resources Applied to VV&A

Criticality

Source: DD(X) Verification, Validation, and Accreditation Overview by Charles Hays of Northrop Grumman Corporation. Presented at NMSO V V & A TWG. Salt Lake City, UT, February 16, 2005.

Figure 1. DD(X) criticality measure

Table 1. Levels of confidence/likelihood of M&S error

Likelihood of error	Confidence level	Description
1	4	Very high confidence based upon extensive documented V&V relevant to intended use.
2	3	High confidence based on face validation by SMEs.
3	2	Moderate confidence based upon previous usage history.
4 (high)	1	Low or unknown level of confidence. M&S appears to have the functionality required but credibility is unknown.

M&S, models & simulations; SME, subject matter experts; V&V, verification & validation.

Table 2. Evidence required to support likelihood of error and accreditation requirements

Likelihood of error	Confidence level	Description
1	4 (high)	Level 3 + Extensive body of documented verification and validation and extensive disciplined M&S development including history of technical and managerial review over time.
2	3	Level 2 + SME face validation relevant to current intended use + Evidence of effective configuration management.
3	2	Level 1 + Usage history + Known V&V history
4 (high)	1	Comparison of M&S requirement derived from intended use with capabilities and limitations of candidate simulation.

M&S, models & simulations; SME, subject matter experts; V&V, verification & validation.

Table 3. An example scheme for "quantifying" likelihood

Likelihood description	Likelihood of occurrence over lifetime of an item	Likelihood of occurrence per number of items
Frequent	Likely to occur frequently	Widely experienced
Probable	Will occur several times in life of item	Will occur frequently
Occasional	Likely to occur some time in life of item	Will occur several times
Remote	Unlikely but possible to occur in life of item	Unlikely but can reasonably be expected to occur
Improbable	So unlikely, it can be assumed occurrence may not be experienced	Unlikely to occur, but possible

Table 4. Critical analysis: importance of the role of M&S

Role level	Definition
4	M&S will be the **only method** employed to make a decision.
3	M&S will be the **primary method,** employed with other non-M&S methods.
2	M&S will be a **secondary method,** employed with other non-M&S methods, and will **provided significant data unavailable through other means.**
1	M&S will be a **supplemental method**, employed with other non-M&S methods, and will **provide supplemental data already available through other means.**

M&S, models & simulations.

Table 5. Critical analysis: importance of decisions

Level	Description
4	Intended use addresses **multiple areas** of significant program risk, key program reviews and test events, key system performance analysis, primary test objectives and test article design, system requirements definition, and/or high software criticality; used to make a technical or managerial decision.
3	Intended use addresses an **area of significant program risk** .
2	Intended use addresses **medium or low program risk**, other program reviews and test events, secondary test objectives and test article design, other system requirements and system performance analysis, and medium or low S/W criticality used to make technical or managerial decisions.
1	Intended use addresses **program objectives or analysis that is not a significant factor** in the technical or managerial decision-making process.

Table 6. Levels of consequences on (cost, schedule & technical) performance if decision is poor

Level	Technical performance	Schedule	Cost
5	Severe degradation in technical performance; cannot meet KPP or key technical/supportability threshold; will jeopardize program success; no workarounds.	Cannot meet key program milestones.	Exceeds APBA threshold.
		Slip ≥months.	>10% of budget.
4	Significant degradation in technical performance or major shortfall in supportability; may jeopardize program success; workarounds may not be available or may have negative consequences.	Program critical path affected, all schedule float associated with key milestone exhausted.	Budget increase or unit production cost increases.
		Slip ≤months	< 10% of budget.
3	Moderate reduction in technical performance or supportability with limited impact on program objectives; workarounds available.	Minor schedule slip, no impact to key milestones.	Budget increase or unit production cost increases.
		Slip <month(s) of critical path. Subsystem slip ≥months.	<5% of budget.
2	Minor reduction in technical performance or supportability; can be tolerated with little or no impact on program; same approach retained.	Additional activities required; able to meet key dates.	Budget increase or unit production cost increases.
		Slip ≤month(s).	<1% of budget.
1	Minimal or no impact.	Minimal or no impact.	Minimal or no impact.

Table 7. Simplified levels of consequence

Consequence level	Definition
High	Major disruption to program. Different approach required. Priority management attention and resource allocation required immediately.
Moderately high	Significant disruption to program. Different approach required. Priority management attention required.
Moderate	Noticeable disruption. Different approach may be required. Additional management attention may be needed.
Low	Minimal impact. Minimum oversight needed to ensure risk remains low.

Table 8. A more advanced scheme for ''quantifying'' consequence (impact) of poor decision

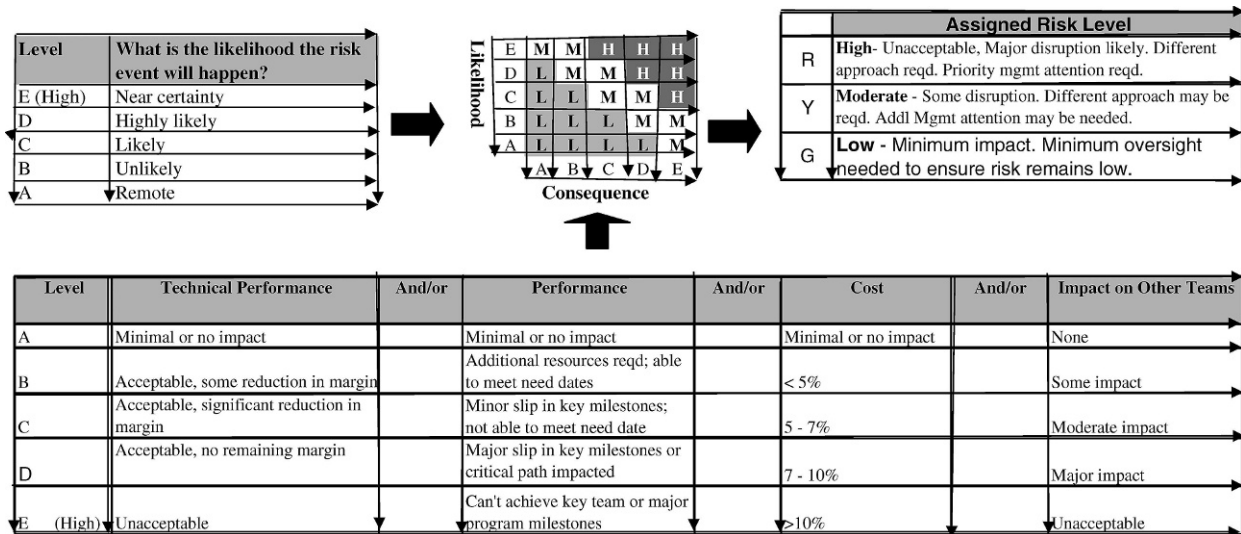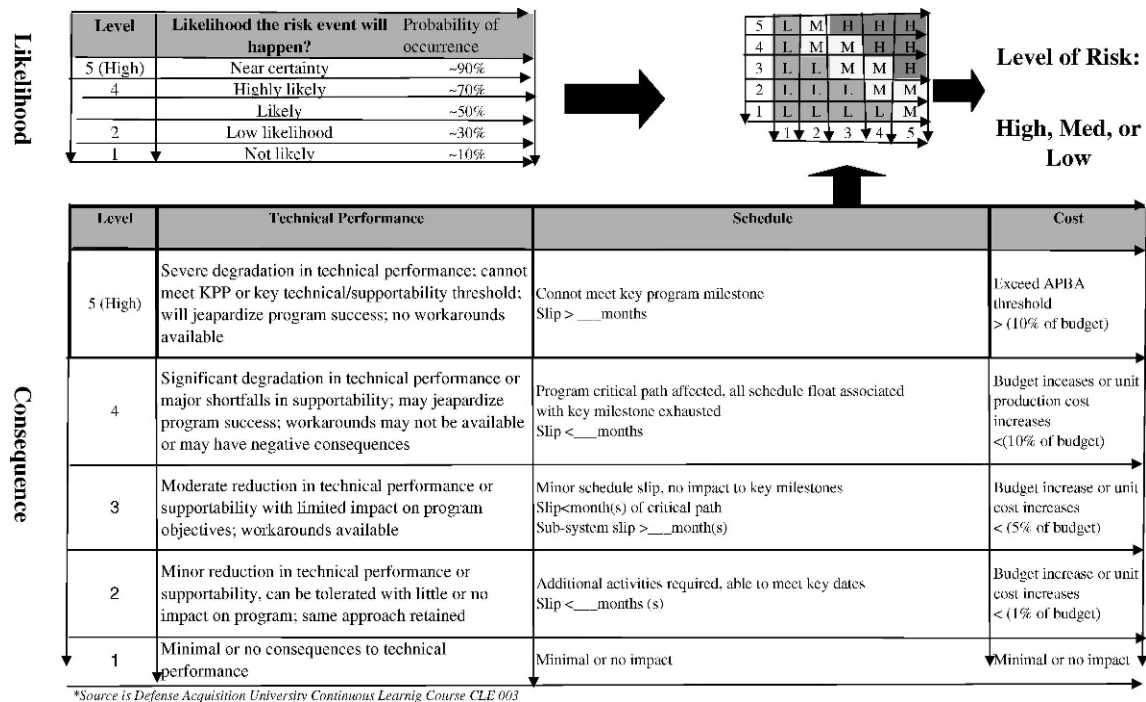|	Impact level			
Impact categories	Catastrophic	Critical	Marginal	Negligible
Personnel safety	Death	Severe injury	Minor injury	<Minor injury
Equipment safety	Major equipment loss broad-scale major damage	Small-scale major damage	Broad-scale minor damage	Small-scale minor damage
Environmental damage	Severe (Chernobyl)	Major (Love Canal)	Minor	Some trivial
Occupational illness	Severe & broad	Severe or broad	Minor and small scale	Minor or small scale
Cost	Loss of program funds: 100% cost growth	Funds reduction: 50% to 100% cost growth	20% to 50% cost growth	<20% cost growth
Schedule	Slip reduces DoD capabilities	Slip causes cost impact	Slip causes internal turmoil	Republish schedules
Political	National or international (Watergate)	Significant (Tailhook)	Embarrassment ($200 hammer)	Local
Operational	Widespread additional combat deaths	Limited additional combat deaths	Moderate additional casualties	Minimal additional casualties

Level	What is the likelihood the risk event will happen?
E (High)	Near certainty
D	Highly likely
C	Likely
B	Unlikely
A	Remote

	Assigned Risk Level
R	**High**- Unacceptable, Major disruption likely. Different approach reqd. Priority mgmt attention reqd.
Y	**Moderate** - Some disruption. Different approach may be reqd. Addl Mgmt attention may be needed.
G	**Low** - Minimum impact. Minimum oversight needed to ensure risk remains low.

Likelihood

E	M	M	H	H	H
D	L	M	M	H	H
C	L	L	M	M	H
B	L	L	L	M	M
A	L	L	L	L	M

A B C D E

Consequence

Level	Technical Performance	And/or	Performance	And/or	Cost	And/or	Impact on Other Teams
A	Minimal or no impact		Minimal or no impact		Minimal or no impact		None
B	Acceptable, some reduction in margin		Additional resources reqd; able to meet need dates		< 5%		Some impact
C	Acceptable, significant reduction in margin		Minor slip in key milestones; not able to meet need date		5 - 7%		Moderate impact
D	Acceptable, no remaining margin		Major slip in key milestones or critical path impacted		7 - 10%		Major impact
E (High)	Unacceptable		Can't achieve key team or major program milestones		>10%		Unacceptable

Figure 2. Example use of standard risk chart

the likelihood of model error is determined by *associating* the level of reliance on the M&S and the level of confidence in the M&S.

- **Rule 3: Value of "consequence of model error" & "likelihood of model error" → risk.** The risk level is finally determined by *associating* the consequence of model error with the likelihood of model error.

We point out to the user that it is in defining these association rules that the tailoring of this process can best be done. Here is where experience, technical skills, thorough understanding of what is being modeled, how it is to be used, the associated costs and schedules, etc., come in to play. Depending on the scenario, the definition of the association rules can become a serious task. It may become necessary to

Likelihood

Level	Likelihood the risk event will happen?	Probability of occurrence
5 (High)	Near certainty	~90%
4	Highly likely	~70%
3	Likely	~50%
2	Low likelihood	~30%
1	Not likely	~10%

5	L	M	H	H	H
4	L	M	M	H	H
3	L	L	M	M	H
2	L	L	L	M	M
1	L	L	L	L	M

1 2 3 4 5

Level of Risk:

High, Med, or Low

Consequence

Level	Technical Performance	Schedule	Cost
5 (High)	Severe degradation in technical performance; cannot meet KPP or key technical/supportability threshold; will jeapardize program success; no workarounds available	Connot meet key program milestone Slip > ___months	Exceed APBA threshold > (10% of budget)
4	Significant degradation in technical performance or major shortfalls in supportability; may jeapardize program success; workarounds may not be available or may have negative consequences	Program critical path affected, all schedule float associated with key milestone exhausted Slip <___months	Budget inceases or unit production cost increases <(10% of budget)
3	Moderate reduction in technical performance or supportability with limited impact on program objectives; workarounds available	Minor schedule slip, no impact to key milestones Slip<month(s) of critical path Sub-system slip >___month(s)	Budget increase or unit cost increases < (5% of budget)
2	Minor reduction in technical performance or supportability, can be tolerated with little or no impact on program; same approach retained	Additional activities required, able to meet key dates Slip <___months (s)	Budget increase or unit cost increases < (1% of budget)
1	Minimal or no consequences to technical performance	Minimal or no impact	Minimal or no impact

Source is Defense Acquisition University Continuous Learnig Course CLE 003

Figure 3. Example of program risk reporting

	Level of Impact			
Probability	Catastrophic	Critical	Marginal	Negligible
Frequent	**High**	**High**	**Medium**	**Low**
Probable	**High**	**High**	**Medium**	**Low**
Occasional	**High**	**Medium**	**Low**	**Low**
Remote	**Medium**	**Low**	**Low**	**Low**
Impossible	**Low**	**Low**	**Low**	**Low**

Indicates risk level

Risk level values are:
- Subjective
- Based on MIL-STD-882C/D (US Military System Safety Standard)
- Tailorable to each problem

Source: MIL-STD 882D, Department of Defense Standard Practice for System Safety, February 10,
Available at www:safetycenter.navy.mil/instructions/osh/milstd882d.pdf

Figure 4. Combination scheme for reduced number of risk levels

brainstorm a bit to determine what is suitable, and it may be necessary to get help from the experts and SMEs.

Risk determination procedure

The standard risk chart *(Figures 2–4)* requires the user to input values of likelihood and consequence. Thus, the use of any combination of the scales given above must result in the values of likelihood and consequence for the final step in the risk determination procedure. The association rules given above are meant to guide the user in defining a combination of schemes that will lead to this final step. A step-by-step recipe is provided with a simplified example to illustrate the process in what follows. No matter which scheme the user chooses, consideration must be given to consequences of the varying nature, including cost, schedule, personnel safety, political (including ridicule), and operational setting. The user must also accommodate all of the ways the model output could be wrong (e.g., erroneously overestimated or underestimated performance for which the consequences might be different in each case).

Conclusion

This article presents a methodology for assessing and prioritizing the risk associated with using M&S to support acquisition decisions. DoD, the Navy, and other Services mandate the use of M&S to support acquisition. Oftentimes limitations of resources make it hard to formally conduct VV&A of these models as needed to establish their credibility. However, using models and simulations whose credibility is in question can involve serious risk. The procedures presented here provide decision-makers with judicious ways for determining the risks associated with using a given M&S, and the extent to which VV&A work will be needed to reduce such risks. The approach presented parallels the formal DoD risk assessment procedure, but with application to the use of M&S as it relates to VV&A. The final output is a ranking of the risk associated with the models as either high, medium, or low risk. With this ranking the decision-maker is now able to decide whether or not to accept the risk or to invest the resources needed to mitigate the risk. An important outcome of this approach is that, through this method, there is the ability to determine what level of authority is needed to serve as the accreditation authority.   ❏

James N. Elele, Ph.D., was born in Nigeria, West Africa, and immigrated to the United States 32 years ago. He earned a Bachelor of Science degree in chemical engineering (1980), a Master of Science degree (1985), and a Ph.D. in Applied Mathematics (1988), all from the University of Arizona, Tucson, Arizona.

He was employed as an engineer at IBM from 1980 to 1981 and at General Electric from 1981 to 1983. He was then employed by the U.S. government at the Electronics Proving Ground, Fort Huachuca, Arizona in 1988. He transferred to Patuxent River, Maryland, in 1997 and has continued there to the present.

Dr. Elele has extensive experience in modeling, simulation, and VV&A. He was part of the team that created the Army's Mobile Subscriber Equipment Performance Model (MSEPAM) and oversaw its VV&A. He served as the data specialist for the Army's Extended Air Defense Test Bed (EADTB) and completed a developmental assignment with the Army Test & Evaluation Command, which was instrumental in defining and introducing the Virtual Proving Ground (VPG).

Dr. Elele currently provides VV&A support to various acquisition programs, while serving as the M&S lead for T&E IPT in support of the development of the Marine Corps Heavy Lift Helicopter (CH–53K) replacement program. He also serves as the acting director of the BVVASO also known as JASA and as an adjunct professor for Strayer University and (formerly for) the Florida Institute of Technology. He has published 24 technical papers in journals & conferences, and he is the owner of two U.S. patents. E–mail: james.elele@navy.mil.

Endnotes

[1]DOD, 2003. Department of Defense Instruction 5000.61, DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A), 12 May 2003.

[2]U.S. Navy, 2004. Department of Navy. "Modeling and Simulation Verification, Validation and Accreditation Implementation Handbook, Volume I: VV&A Framework." March 30, 2004.

[3]DMSO, 2000. Defense Modeling and Simulation Office (DMSO). Risk and Its Impact on VV&A. DMSO VV&A Recommended Practices Guide Special Topic Paper, November 30, 2000.

[4]BVVASO, (1997). VV&A from A to Z, Joint Accreditation Support Activity The Battlespace Verification, Validation, and Accreditation Support Office (BVVASO)/JASA publication on practical Work Breakdown on How to conduct VV&A.

[5]DMSO, 1996. Department of Defense "Verification, Validation, and Accreditation (VV&A) Recommended Practices Guide," November 1996. Available at: www.vva.dmso.mil.

References

Commander Operational Test Force, US Navy. "Navy Modeling and Simulation (M&S) Management." Chief of Naval Operations Instruction 5200.34, May 28, 2002.

Commander Operational Test Force, US Navy. "Use of Modeling and Simulation (M&S) in Operational Testing." COMOPTEVFORINST 5000.1A, September 5, 1995.

Commander Operational Test Force, US Navy. "Use of Modeling and Simulation (M&S) in Operational Testing." COMOPTEVFORINST 5000.1, September 9, 2004.

Commander Operational Test Force, US Navy. "COMOPTEVFOR Modeling and Simulation Accreditation Documentation." COMOPTEVFOR Policy and Information Notice 99-01, May 13, 1999.

Commander Operational Test Force, US Navy. "Accreditation of Modeling and Simulation in Support of Operational Test and Evaluation." COMOPTEVFOR Letter 3980 Ser 00T/312, April 29, 1998.

US Army. "Management Verification, Validation, and Accreditation of Army Models and Simulations." Department of the Army Pamphlet 5–11. September 30, 1999.

US Army. "Management Simulation Support Planning and Plans." Department of the Army Pamphlet 5–12, March 2, 2005.

US Department of Defense. "DoD Modeling and Simulation (M&S) Management." Department of Defense Directive 5000.59, January 4, 1994.

US Department of Defense. "DoD Modeling and Simulation (M&S) Management." DoD Directive 5000.59, January 4, 1994.

US Department of Defense. "DoD Glossary of M&S Terms." DoD 5000.59-M, December 1997.

US Department of Defense. "Modeling and Simulation (M&S) Master Plan." DoD 5000.59-P, October 1995.

US Department of Defense. "DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A)." DoD INST 5000.61, April 29, 1996.

US Department of Defense. "Introduction to VV&A Standards and Practices for DoD M&S" (Draft), September 30, 1995.

Elele, J. E. 2007. "Assessing and Prioritizing the Risks Associated with the Level of Verification, Validation and Accreditation." *Proceedings of the NDIA 10th Annual Systems Engineering Conference.* San Diego, California, September 2007.

Elele, J. E. 2007. "A Methodology for Assessing and Prioritizing the Risks Associated with the Level of Verification, Validation and Accreditation of Models and Simulations." *Proceedings Annual ITEA Conference.* Kauai, Hawaii, November 2007.

Kilikauskas, M. and Hall, D. H. YEAR "The Use of M&S VV&A as a Risk Mitigation Strategy in Defense Acquisition," *JDMS*, 2, (4), p. 215.

Kilikauskas, M., Hall, D. H. and Weeks, J. 2003. "Comparison and Analysis of DoD and Service Policy and Guidance on Modeling and Simulation (M&S) Verification, Validation and Accreditation." Joint Aircraft Survivability Program Office Task M-7-03, December 2003.

"Department of Defense Standard Practice for System Safety." MIL-STD 882D, February 10, 2000, www.safetycenter.navy.mil/instructions/osh/milstd882d.pdf

MIL-STD-882 Rev D of February 10, 2000 is current issue, Rev E dated 1 February 1, 2000 is out in draft.

"Operational Risk Management Policy." NAVTESTWINGLANT/NAVTESTWINGPAC INSTRUCTION 3500-1, December 11, 2002.

NMSO. "Modeling and Simulation, Verification, Validation, and Accreditation Implementation Handbook, Volume I, VV&A Framework."

US Navy. "Navy Modeling and Simulation Master Plan." February 21, 1997.

US Department of Defense. "Risk Assessment and Its Impact on VV&A." DMSO RPG, November 30, 2000.

Defense Acquisition University. "Risk Management Guide for DoD Acquisition, 6th Edition," August 2006, Defense Acquisition University.

US Navy. "Department of the Navy Modeling and Simulation Program." SECNAVINST 5200.38, October 18, 1994.

US Navy. "VV&A of Models and Simulations." SECNAVINST 5200.40, April 19, 1999.

US Navy. "Department of the Navy Modeling and Simulation Management." Secretary of the Navy Instruction 5200.38A, February 28, 2002.

US Navy. "Verification, Validation, and Accreditation (VV&A) of Models and Simulations." Secretary of the Navy Instruction 5200.40A (Draft), November 3, 2003.

U.S. Air Force. "Operations Support: Verification, Validation and Accreditation (VV&A)." Air Force Instruction AFI 16-1001. June 1, 1996.

U.S. Department of Defense. "DoD Modeling and Simulation (M&S) Management." DoD Directive 5000.59. (Administrative Change 1 to the original document dated January 4, 1994). January 20, 1998.

U.S. Department of Defense. "Operation of the Defense Acquisition System." DoD Instruction 5000.2. May 12, 2003.

U.S. Department of Defense. "DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A)." DoD Instruction 5000.61. May 13, 2003.

U.S. Navy. "Verification, Validation and Accreditation (VV&A) of Models and Simulations." Secretary of the Navy Draft Instruction SECNAVINST 5200.40A (in coordination), January 20, 2005.

"Verification, Validation and Accreditation of Joint Models and Simulations." Joint Staff Instruction JSI 8104.01A, January 26, 2004.

Acknowledgments